
(19) KOREAN INTELLECTUAL PROPERTY OFFICE

KOREAN PATENT ABSTRACTS

(11)Publication number: 1020020068585 A
(43)Date of publication of application: 28.08.2002

(21)Application number: 1020010008666
(22)Date of filing: 21.02.2001

(71)Applicant: SAMSUNG ELECTRONICS CO., LTD.
(72)Inventor: KOO, JEONG HOE

(51)Int. Cl. G10L 19/12

(54) METHOD OF REDUCING AMOUNT OF CALCULATION REQUIRED FOR PITCH SEARCH IN VOCODER

(57) Abstract:

PURPOSE: A method of reducing the amount of calculation required for pitch search in a vocoder is provided to decrease the amount of calculation needed for pitch search by restricting a section that requires correlation calculation for pitch search.

CONSTITUTION: When a voice signal is inputted (202), the voice signal is passes through an analysis filter and a synthesis filter(204,206). The voice signal that passed through the filters passes through a cubing filter(208). The cubing filter emphasizes a high-amplitude portion of the voice signal but suppresses a low-amplitude portion of the voice signal. The voice signal that passed through the cubing filter passes through a clipping filter that removes a portion of the voice signal lower than a threshold value(210). A high-amplitude region of the voice signal that passed through the clipping filter is determined(212), and pitch delay and pitch gain are obtained.

© KIPO 2003

Legal Status

Date of request for an examination (20010221)

Final disposal of an application (registration)

Date of final disposal of an application (20030409)

Patent registration number (1003835890000)

Date of registration (20030429)

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。 Int. Cl. ⁷
G10L 19/12

(11) 공개번호 특2002-0068585
(43) 공개일자 2002년08월28일

(21) 출원번호 10-2001-0008666
(22) 출원일자 2001년02월21일

(71) 출원인 삼성전자 주식회사
경기 수원시 팔달구 매탄3동 416
(72) 발명자 구정희
서울특별시성북구정릉2동206-120
(74) 대리인 이견주

심사청구 : 있음

(54) 보코더에서 피치 검색에 필요한 계산량을 감소하는 방법

요약

본 발명은 보코더에서 필요한 계산량을 감소한 피치 검색 방법에 있어서, 음성 신호를 입력되면 분석 필터 및 합성 필터를 통과시키는 제1 단계와, 상기 제1 단계의 결과인 음성 신호를 고진폭 부분을 강화시켜주고, 저진폭 부분을 억제시켜주는 커빙 필터를 통과시키는 제2 단계와, 상기 제2 단계의 결과인 음성 신호를 음성 신호 중에서 임계값보다 작은 부분을 모두 제거하여 주는 클리핑 필터를 통과시키는 제3 단계와, 상기 제3 단계의 결과인 음성 신호에서 고진폭 영역을 결정하고, 피치 지연과 피치 이득을 구하는 제4 단계를 포함하는 것을 특징으로 한다.

대표도
도 4

색인어
피치 검색, 커빙 함수, 클리핑 함수

명세서

도면의 간단한 설명

도 1은 종래 피치 검색의 흐름도를 나타내는 도면,

도 2는 본 발명의 실시예에 따른 CELP 압축기의 디코더 구성을 도시한 도면

도 3a는 클리핑(Clipping) 함수를 나타내는 도면,

도 3b는 커빙(Cubing) 함수를 나타내는 도면,

도 3c는 입력 음성 신호를 나타내는 도면,

도 3d는 도 3c의 입력 음성 신호가 클리핑 필터를 통과한 신호를 나타내는 도면,

도 3e는 도 3c의 입력 음성 신호가 클리핑 필터 및 커빙 필터를 통과한 신호를 나타내는 도면,

도 4는 본 발명에 따른 피치 검색 방법을 나타내는 제어 흐름도.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 보코더에 관한 것으로, 특히 보코더에서의 피치 검색에 관한 것이다.

디지털 통신은 전달하고자 하는 정보를 디지털 신호로 부호화하여 전달하는 통신 방식이다. 따라서 디지털 음성 통신을 위하여 아날로그 신호인 음성 신호를 디지털 부호로 변환하여야 하며, 이때 원하는 통화 음질을 보장하면서 전달하는 디지털 데이터의 양을 줄여 보다 효율적인 통신이 되도록 하여야 한다. 이를 위하여 음성 신호를 압축시켜 전송하는 기술을 이용하며, 아날로그 음성 신호를 적은 양의 디지털 신호로 압축하는 장치를 엔코더(encoder), 압축된 신호로부터 원 음성 신호를 재생하는 장치를 디코더(Decoder)라 하고, 두 부분을 합하여 음성 압축기(Vocoder)라 한다. 디지털 음성 통신은 항상 음성 압축기를 거쳐 이루어지며 음성 압축기에 의하여 통화 음질과 데이터 양이 결정되므로 음성 압축기는 디지털 통신의 품질과 용량을 결정하는 매우 중요한 부분이다.

현재 여러 국제 통신 기구들에 의하여 많은 표준 압축기들이 개발되어 다양한 종류의 통신 시스템에 적용되어 서비스 중에 있다. 그러나 이들 표준 음성 압축기들은 각각 그 사용 환경에 따라 설계상의 제약점이 서로 틀리고 이로 인하여 압축 방법에 큰 차이를 가진다. 따라서 음성 압축기의 성능을 직접적으로 비교하는 것은 의미가 없고 주어진 사용 환경과 조건에서의 성능을 살펴보아야 하며, 각 음성 압축기를 원래 의도된 사용 환경과 다른 환경에서 동작시키면 예상된 성능을 보이지 못하거나 또는 Bit를 비효율적으로 사용하게 되는 등의 문제점이 발생한다. 일례로, G.726 32kbps AD PCM[1]은 유선망에서는 매우 우수한 음질을 가지지만 이를 무선 통신에 사용하면 데이터 전송 에러에 의하여 음질이 급격히 떨어져 사용이 힘들어 지며, 또한 CDMA에 사용되는 IS-96A QCELP 음성 압축기[2]를 데이터 전송 에러가 없는 유선망에서 사용하면 Bit의 사용이 비효율적이 된다(QCELP는 전송 에러를 막기 위한 목적으로 11 Bit를 사용한다). 따라서 새로운 통신 서비스가 개발되고 이를 위한 새로운 통신 환경이 정해질 경우 기존의 다른 환경용으로 설계된 음성 압축기를 그대로 사용하면 음질 저하 또는 자원의 비효율적 사용으로 인하여 통신 시스템 전체의 품질이 저하되고 이로 인하여 통신 서비스의 경쟁력을 잃게 된다. 이를 해결하기 위하여 새롭게 정의된 환경에 가장 적합한 새로운 음성 압축기의 개발이 필수적이며 이를 통하여 경쟁력 있는 통신 서비스의 제공이 가능하여 진다.

일반적으로 음성 신호를 코딩하기 위한 보코더 기술에는 3가지 방법, 즉 파형 코딩 방식, 소스 코딩 방식 및 하이브리드 방식이 있다. 이들 방식 중 하이브리드 방식이 보코더용으로 가장 바람직한 방법으로 사용되고 있다. 이 하이브리드 방

법은 선형 예측 분석을 이용함으로써 발성 추적 필터(vocal tract filter)를 모델링하고 특히 이들을 코딩한 후의 나머지 잔여 신호를 전송한다. 하이브리드 방법에는 3가지의 예측 방법, 즉 RELP(Residual Excited Linear Prediction), VELP(Voice Excited Linear Prediction) 및 CELP(Code Excited Linear Prediction)이 있다.

이들 예측 방법 중, 비트 레이트(bit rate)를 기준으로 가장 우수한 성능을 나타내고 있는 것으로 알려진 CELP (Code book Excited Linear Prediction) 계열의 보코더인 EVRC(Enhanced Variable Rate Coder)와 QCELP(Qualcomm Codebook Excited Linear Prediction) 보코더가 각각 셀룰러, PCS와 같은 무선 음성 통신 시스템에 사용되고 있다.

통신 시스템에서, 송신단에서 전송한 사람의 음성 데이터를 수신단에서 복조했을 때, 제한된 자원(resource)인 주파수 효율성을 높이면서 음성통화를 위한 통화품질을 최대한 양호하게 하기 위한 CELP 알고리즘에는 음성 생성 모델에 근거한 AbS(Analysis by Synthesis) 방법에 근거해서 선형 예측 모델을 위한 LPC(Linear Prediction Coefficient)계산, 피치 검색(pitch search), 코드북 검색(codebook search) 등의 서브 블록(sub-block)들을 포함하고 있다. 이러한 CELP의 서브 블록들 중, 피치 검색은 음성 신호의 자기 상관 관계에 해당하는 피치 주기에 관한 정보를 얻어내는 과정이다.

일반적으로 보코더에서는, 시간에 따라 약간씩 변화하는 음성 신호의 특성을 고려하여, 5msec~10msec 정도의 구간에 대해서 피치검색을 수행한다. 입력 음성 신호를 8kHz로 샘플링 했을 경우, 이러한 피치 검색 구간에 대해서는 시불변(time invariant)이라는 가정 하에, 피치 지연(L)과 피치 이득(b)을 구하는데, 이때는 보통 피치 지연을 20 ~ 147 정도로 제한한다.

도 1은 이러한 피치 검색의 흐름도를 나타내는 도면이다.

보코더는 먼저 단계 102 및 단계 104에서 입력 신호를 수신하여 분석 필터(analysis filter)를 통과시킨다. 이어서 보코더는 단계 106에서 분석 필터를 통과한 잔여 신호에 대해서 합성 필터(synthesis filter)의 영입력 응답을 제거한 신호 $x(n)$ 를 발생하고, 단계 108에서 신호 $x(n)$ 와 이전 프레임으로부터 구해지는 적응 코드북(adaptive codebook) $y_L(n)$ 에 대해서, 위에서 제한한 피치 지연(L)에 대해서 피치 검색을 통해 다음과 같이 에러(E_L)를 최소로 만드는 피치 지연(L)과 피치 이득(b)을 구하게 된다. 이는 다음 수학적 식 1과 같다.

수학적 식 1

$$E_L = \sum_{n=0}^{N-1} \{x(n) - by_L(n)\}^2$$

여기에서 N은 피치 검색을 위한 서브 프레임의 음성 신호의 개수이다.

이때, E_L 에 대한 최소값은 다음 수학적 식 2를 최소로 하는 L, b를 구하면 된다.

수학적 식 2

$$\begin{aligned} P &= -2bE_{xy} + b^2E_{yy} \\ E_L &= \sum_{n=0}^{N-1} x(n)^2 \\ E_L &= \sum_{n=0}^{N-1} x(n) \cdot y_L(n) \\ E_L &= \sum_{n=0}^{N-1} y_L(n)^2 \end{aligned}$$

여기에서, E_L 에 대한 최소값을 구하는데 있어서 E_{xx} 는 상수이므로 고려하지 않는다. 먼저 P 는 b 에 대한 아래로 볼록인 2차 함수이므로 P 를 최소로 하는 b 는 다음 수학식 4와 같이 구한다.

수학식 3

$$\frac{\partial P}{\partial b} = -2E_{xy} + 2bE_{yy} = 0$$

$$b = \frac{E_{xy}}{E_{yy}}$$

그리고, 수학식 3을 수학식 2에 대입하면, 다음 수학식 4가 도출된다.

수학식 4

$$P_{\min} = -2 \cdot \frac{E_{xy}}{E_{yy}} \cdot E_{xy} + \left(\frac{E_{xy}}{E_{yy}}\right)^2 \cdot E_{yy} = -\frac{E_{xy}^2}{E_{yy}}$$

발명이 이루고자 하는 기술적 과제

그런데, 전술한 종래의 보코더에서의 피치 검색은 P_{\min} 에 대한 L , b 를 구하는 과정으로서, 자기 상관(auto-correlation) 및 상호 상관(cross-correlation)에 많은 계산량이 소요된다. 결국, 피치 검색은 CELP 전체 코더(부호화기)의 계산량에서 25% 정도를 차지할 정도로 방대한 계산량을 필요로 한다. 따라서, 음질 저하를 최소화시키면서 이에 필요한 계산량을 줄이는 것은 저전력을 필요로 하는 무선 통신 장치에의 적용에 필수 불가결하다.

따라서, 본 발명의 목적은 피치 검색에 필요한 계산량을 줄일 수 있는 방법을 제공함에 있다.

본 발명의 다른 목적은 피치 검색에 필요한 계산량을 줄임으로써 피치 검색에 소요되는 시간을 감소할 수 있는 방법을 제공함에 있다.

발명의 구성 및 작용

상기 목적을 달성하기 위한 본 발명은 보코더에서 필요한 계산량을 감소한 피치 검색 방법에 있어서, 음성 신호를 입력 되면 분석 필터 및 합성 필터를 통과시키는 제1 단계와, 상기 제1 단계의 결과인 음성 신호를 고진폭 부분을 강화시켜 주고, 저진폭 부분을 억제시켜 주는 커빙 필터를 통과시키는 제2 단계와, 상기 제2 단계의 결과인 음성 신호를 음성 신호 중에서 임계값보다 작은 부분을 모두 제거하여 주는 클리핑 필터를 통과시키는 제3 단계와, 상기 제3 단계의 결과인 음성 신호에서 고진폭 영역을 결정하고, 피치 지연과 피치 이득을 구하는 제4 단계를 포함한다.

이하 본 발명에 따른 바람직한 실시 예를 첨부한 도면을 참조하여 상세히 설명한다. 그리고 본 발명을 설명함에 있어서, 관련된 공지 기능 혹은 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

도 2은 본 발명의 실시예에 따른 CELP 압축기의 디코더 구성을 도시한 도면이다. 이 CELP 압축기의 디코더는 Analysis-by-Synthesis 방법을 이용하여 예측 오차 신호를 코드북(Codebook)을 이용하여 코드(Code)화하는 방법이다.

도 2에서, 도면 부호 12는 피치 필터를 나타내고 도면 부호 22는 코드북(codebook)을 나타낸다. 도면 부호 14, 24는 믹서, 도면 부호 16은 가산기를 나타낸다. 또, 도면 부호 18은 LPC 필터, 도면 부호 20은 포스트필터를 나타낸다.

도 2의 디코더(30)는 엔코더(신호 압축기)로부터 전달된 코드북 인덱스(Codebook Index)와 피치 래그(Pitch Lag) 및 각각의 게인(G_p , G_c)을 이용하여 여기(Excitation) 신호를 만들고 이 신호를 LPC 필터링(Liner Prediction Coding Filtering)하여 음성을 합성한다. 마지막으로 합성된 음성 신호의 음질을 향상시키기 위하여 포스트필터(Postfilter)를 이용하여 최종 음성 신호를 출력시킨다.

도 3a는 클리핑(Cliping) 함수를 도 3b는 커빙(Cubing) 함수를 나타내는 도면이고, 도 3c는 입력 음성 신호를 나타내는 도면이다. 또한, 도 3d는 도 3c의 입력 음성 신호가 클리핑 필터를 통과한 신호를 나타내는 도면이고, 도 3e는 도 3c의 입력 음성 신호가 클리핑 필터 및 커빙 필터를 통과한 신호를 나타내는 도면이다.

본 발명에서 저진폭(low amplitude) 부분을 적절히 제거하기 위해 전처리기인 커빙 함수(도 3a)와 클리핑 함수(도 3b)는 다음 수학적식 5 및 수학적식 6과 같은 전달 함수를 갖는다.

수학적식 5

$$Z(x) = x^3$$

수학적식 6

$$C(x) = \begin{cases} x-T, & x > T \\ 0, & |x| \leq T \\ x+T, & x < -T \end{cases}$$

상기 수학적식 5의 커빙 함수는 음성 신호(도 3c)의 고진폭(high amplitude) 부분을 강화시켜주고, 저진폭 부분을 억제시켜 주는 역할을 한다. 상기 수학적식 6의 클리핑 함수는 음성 신호 중에서 임계(threshold)값 T 보다 작은 부분을 모두 제거하는 역할을 한다. 따라서 이때 임계값 T 는 음성 신호의 클리핑 정도를 결정하는 중요한 역할을 하게 되고, 이는 전체 시스템의 성능에 큰 영향을 미치게 되므로 설정에 많은 주의를 요한다. 즉, T 를 너무 큰 값으로 설정하면 입력 신호를 너무 과도하게 클리핑하게 되고, 너무 작은 값으로 설정하면 클리핑 효과를 얻을 수 없다.

본 발명에서는 커빙 함수와 클리핑 함수를 연관(concatenation)시킴으로써, 입력 음성 신호의 변동에 보다 안정적으로 임계값 T 를 설정할 수 있도록 구성한다. 즉, 커빙 함수를 통해 저진폭 부분과 고진폭 부분의 차이를 보다 극대화시킴으로써 클리핑 함수의 임계값 T 의 설정에 필요한 마진(margin)을 확보할 수 있다. 즉, T 값 설정에 따른 시스템의 안정도를 높여줄 수 있다. T 값은 통상 입력 신호 레벨에 따라 경험적으로 정해줄 수 있으며, 본 발명에서는 T 를 입력 음성 신호의 최대 크기의 30%로 설정한다.

도 2d는 입력 음성 신호를 클리핑 함수만 적용한 경우를 나타내고, 도 2e는 입력 음성 신호를 커빙 함수 및 클리핑 함수를 통과시킨 후의 결과를 나타낸다.

도 2e에 도시된 바와 같이, 음성 신호를 전처리단을 통과시킴으로써 고진폭 영역 $L1-L2$, $L3-L4$, $L5-L6$ 이 결정되면 이 영역에 대해서만 상관 계산을 수행하면 되므로, 피치 검색에 필요한 계산량을 줄일 수 있다. 따라서, 수학적식 7의 결과가 얻어진다.

수학식 7

$$E_y = \sum_{n=L_1}^{L_1-1} x(n) \cdot y_L(n) + \sum_{n=L_2}^{L_2-1} x(n) \cdot y_L(n) + \sum_{n=L_3}^{L_3-1} x(n) \cdot y_L(n)$$

$$E_y = \sum_{n=L_1}^{L_1-1} y_L(n)^2 + \sum_{n=L_2}^{L_2-1} y_L(n)^2 + \sum_{n=L_3}^{L_3-1} y_L(n)^2$$

도 4는 본 발명에 따른 피치 검색 방법을 나타내는 흐름도를 나타내는 도면이다. 도 4의 흐름도에서 단계 202, 204 및 206은 도 1의 단계 102, 104, 106과 동일하며 도 1의 단계 108은 도 4의 단계 214와 동일하며 그에 따라 그 설명을 생략한다.

본 발명에 따른 피치 검색의 흐름도는 종래의 피치 검색의 흐름도에서 단계 208, 단계 210 및 단계 212를 부가한 것이다. 먼저 도 2에 도시된 디코더는 단계 202 및 단계 204에서 입력 신호를 수신하여 분석 필터(analysis filter)를 통과시킨다. 이어서 보코더는 단계 106에서 분석 필터를 통과한 잔여 신호에 대해서 합성 필터(synthesis filter)의 영입력 응답을 제거한 신호 $x(n)$ 를 발생한다. 그런 다음, 보코더는 단계 208 및 210에서 신호 $x(n)$ 를 커빙 필터 및 클리핑 필터를 통과시킨다. 전술한 바와 같이, 도 3e에 도시된 신호가 커빙 필터 및 클리핑 필터를 모두 통과하여 얻어진 결과이다. 보코더는 단계 212에서 이 결과 신호에서 고진폭 영역을 결정하고, 단계 214에서 신호 $x(n)$ 과 이전 프레임으로부터 구해지는 적응 코드북(adaptive codebook) $y_L(n)$ 에 대해서, 위에서 제한한 피치 지연(L)에 대해서 피치 검색을 통해 수학식 7의 에러(E_L)를 최소로 만드는 피치 지연(L)과 피치 이득(b)을 구하게 된다.

이와 같이, 본 발명에 따른 보코더는 피치 검색에 필요한 수학식 2의 계산량을 수학식 7과 같이 줄임으로써 전체 보코더 계산량의 25% 정도를 차지하는 피치 검색에 필요한 계산량의 상당 부분을 줄일 수 있다. 따라서 이러한 계산량의 감소는 디지털 신호 처리로 구현시 전력 소비를 줄일 수 있다.

발명의 효과

본 발명에서는, 피치 검색을 위한 상관 계산이 필요한 구간을 제한함으로써 전체 보코더 계산량의 25% 정도를 차지하는 피치 검색에 필요한 계산량의 상당 부분을 줄일 수 있다.

(57) 청구의 범위

청구항 1.

보코더에서 필요한 계산량을 감소한 피치 검색 방법에 있어서,

음성 신호를 입력되면 분석 필터 및 합성 필터를 통과시키는 제1 단계와,

상기 제1 단계의 결과인 음성 신호를 고진폭 부분을 강화시켜주고, 저진폭 부분을 억제시켜 주는 커빙 필터를 통과시키는 제2 단계와,

상기 제2 단계의 결과인 음성 신호를 음성 신호 중에서 임계값보다 작은 부분을 모두 제거하여 주는 클리핑 필터를 통과시키는 제3 단계와,

상기 제3 단계의 결과인 음성 신호에서 고진폭 영역을 결정하고, 피치 지연과 피치 이득을 구하는 제4 단계를 포함하는 것을 특징으로 하는 피치 검색 방법.

청구항 2.

제1항에 있어서, 상기 제4 단계의 피치 지연 및 피치 이득은 하기 수학식에 의해 구해지는 것을 특징으로 하는 피치 검색 방법.

$$E_{xy} = \sum_{n=L_2}^{L_2-1} x(n) \cdot y_L(n) + \sum_{n=L_2}^{L_2-1} x(n) \cdot y_L(n) + \sum_{n=L_2}^{L_2-1} x(n) \cdot y_L(n)$$

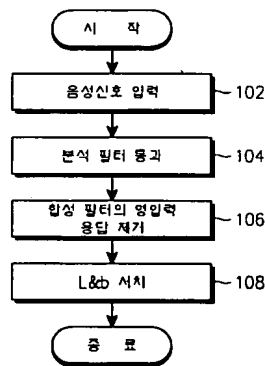
$$E_{yy} = \sum_{n=L_2}^{L_2-1} y_L(n)^2 + \sum_{n=L_2}^{L_2-1} y_L(n)^2 + \sum_{n=L_2}^{L_2-1} y_L(n)^2$$

청구항 3.

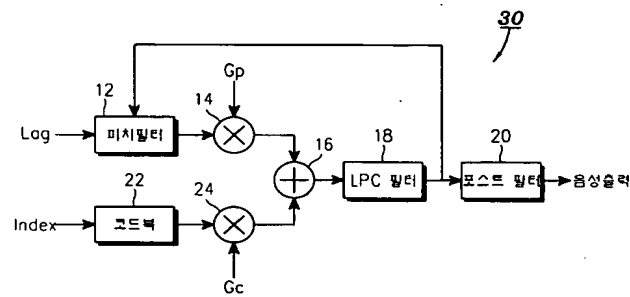
제1항에 있어서, 상기 제4 단계는 상기 고진폭 영역에 대해서만 피치 검색을 위한 상관 계산을 수행함으로써 계산량을 감소하는 것을 특징으로 하는 피치 검색 방법.

도면

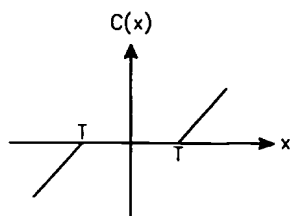
도면 1



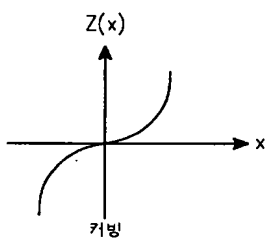
도면 2



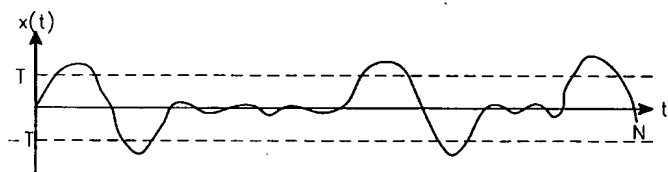
도면 3a



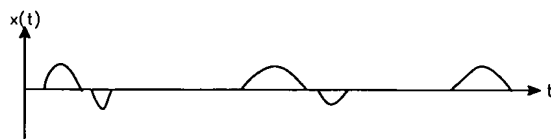
도면 3b



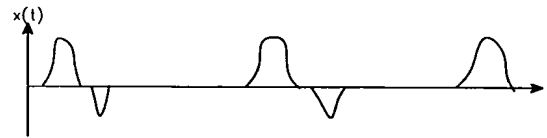
도면 3c



도면 3d



도면 3e



도면 4

